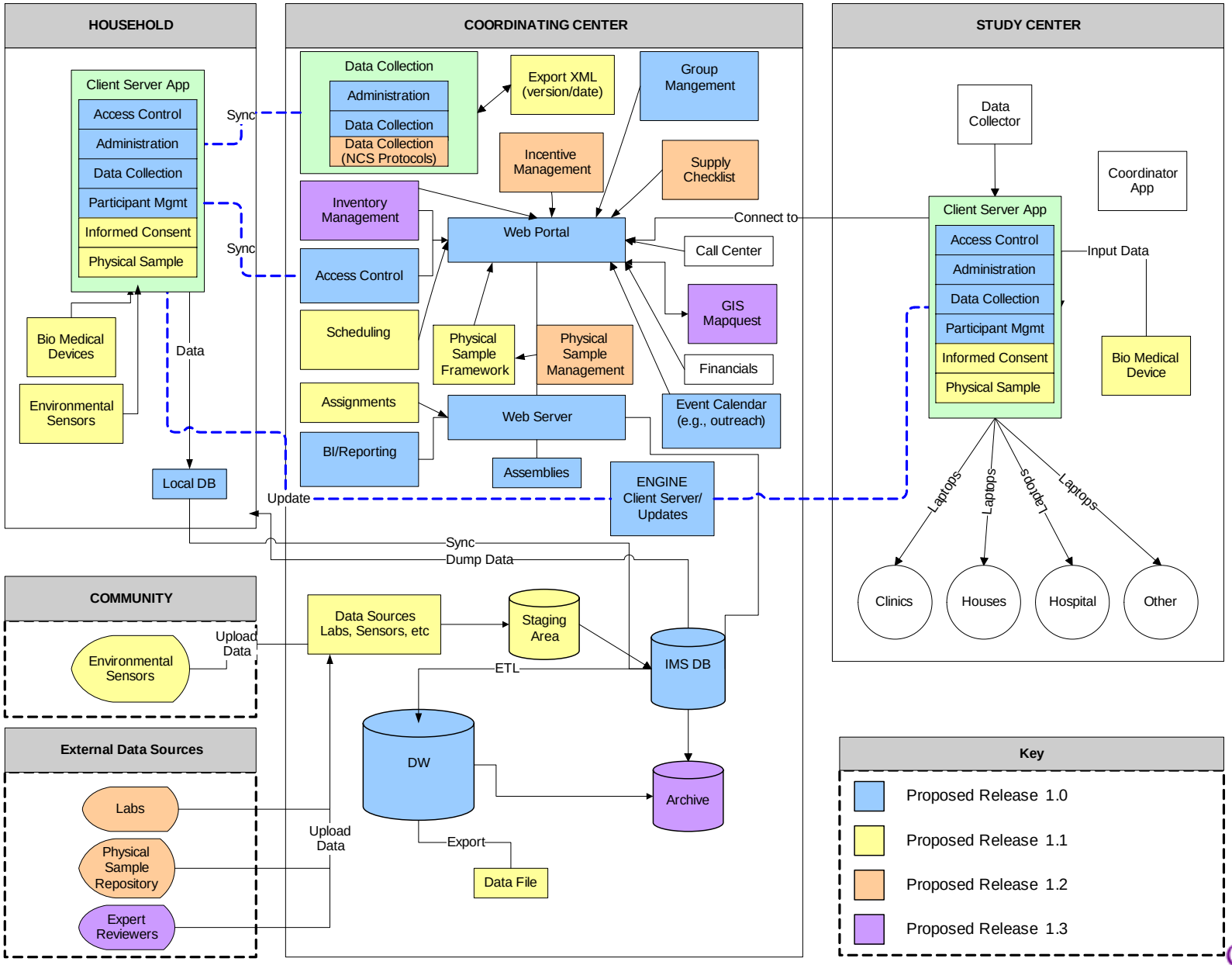
National Children's Study Information
DRAFT 03/15/2005

This presentation was prepared by Booz Allen Hamilton.

THE NATIONAL CHILDREN'S STUDY
HEALTH GROWTH ENVIRONMENT

Current Thinking on National Children's Study Information Management System Logical Architecture - DRAFT



03-15-2005

THE NATIONAL CHILDREN'S STUDY
HEALTH GROWTH ENVIRONMENT

National Children's Study Information Management System Logical Architecture Description – DRAFT 03/15/2005

The National Children's Study architecture must support access from multiple locations. The diagram presents a **draft** logical architecture for the National Children's Study Information Management System (NCS IMS). It depicts the current thinking (03/15/2005) of the NCS IMS team regarding an architecture to meet the design parameters and provide a core architecture that can be scaled to meet the goals of the entire National Children's Study. The following briefly describes the components of the architecture:

- **Coordinating Center:** The Coordinating Center houses the central repository for all data associated with the NCS IMS. This reflects a centralized model where all locations communicate with the Coordinating Center. Centralizing computing resources better enables economies of scale (e.g., higher storage utilization, higher staff utilization) and lowers maintenance costs. Using a centralized model allows services such as data scrubbing and transformation to be shared by all and therefore does not require each site to maintain the tools and associated computing equipment to perform these functions. The central site will receive all data, scrub the data to de-identify participants, and provide regional-specific data back to the Vanguard sites for use by the local study scientists.

- **Study Centers:** The Study Centers have a full time connection to the Coordinating Center, but must have the ability to continue to collect data should the connection to the Coordinating Center fail and then send the collected information to the coordinating center once the connection is re-established. Devices at the Study Centers could be stationary (desktops) or mobile (laptops or tablet PCs). The Study Center user will have a client-server application that will store data locally when not connected to the Coordinating Center, and will also have the ability to interface with biometric devices. The Study Center devices will have the ability to automatically update to the latest version of the client server software via the Client Server Engine Update. This reduces the need to have staff available to manually install software updates.

THE NATIONAL CHILDREN'S STUDY
HEALTH GROWTH ENVIRONMENT

National Children's Study Information Management System Logical Architecture Description – DRAFT 03/15/2005 (continued)

- **Household:** The Household represents the houses at which the Data Collectors will be gathering information. The Data Collectors will be working from mobile devices, such as a laptop or a tablet PC, and will have to have the ability to work in a disconnected mode (e.g. without a live connection to the Coordinating Center). The Household device will have a client-server application that will store data locally. The client-server application will also have the ability to interface with biometric devices and environmental sensor devices at the household. When connected to the Coordinating Center, the Household devices will have the ability to automatically update to the latest version of the client server software via the Client Server Engine, to download all necessary data and questionnaire templates for data collection, and to upload all data collected to the Coordinating Center for storage.

- **Community:** The Community represents the potential for a number of environmental sensors to be located throughout communities participating in the National Children's Study. The data from these sensors will be uploaded to the Coordinating Center on a defined interval.

- **External Data Sources:** External Data Sources are a number of interfaces with laboratories, physical sample repositories, expert reviewers and other external data sources.

THE NATIONAL CHILDREN'S STUDY
HEALTH GROWTH ENVIRONMENT

National Children's Study Information Management System Logical Architecture Description – DRAFT 03/15/2005 (continued)

- **Storage:** The storage system is the core of the NCS IMS. It must be highly available, reliable and extensible enough to handle a wide variety of data types. The storage system must also provide the scalability required for a long-term project such as the National Children's Study. Although initially sized for the initial task, the storage system must be able to scale by several orders of magnitude during the life cycle of the study. The storage system must provide for the storage of various types of data, including: raw data that represents data as it was submitted, normalized data that has been transformed, and de-normalized data for analysis and reporting. These different types of data have a variety of storage requirements based on the frequency of access.

- **Security and Privacy**: The NCS IMS security architecture focuses on the technical controls used to protect the system. These controls are deployed across the input, interface, server and storage components of the NCS system. Technical controls combined with management controls, operational controls and security policies and procedures will mitigate the NCS IMS risks, protect participant privacy, and support system availability, reliability, and retention.

THE NATIONAL CHILDREN'S STUDY
HEALTH GROWTH ENVIRONMENT

Acronyms

- NCS: National Children's Study
- NCS IMS: National Children's Study Information Management System
- XML: eXtensible Markup Language
- GIS: Geographic Information System
- Sync: Synchronize
- IMS DB: Information Management System Database
- DW: Data Warehouse
- DB: Database
- BI: Business Intelligence
- ETL: Extract, Transform, Load
- Mgmt: Management

THE NATIONAL CHILDREN'S STUDY
HEALTH GROWTH ENVIRONMENT